

2000

Effects on stutter production during amplification of tetranucleotide STR loci

Cynthia Renae Hall
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Hall, Cynthia Renae, "Effects on stutter production during amplification of tetranucleotide STR loci" (2000). *Master's Theses*. 2086.
DOI: <https://doi.org/10.31979/etd.kjet-2vzp>
https://scholarworks.sjsu.edu/etd_theses/2086

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

EFFECTS ON STUTTER PRODUCTION DURING AMPLIFICATION OF
TETRANUCLEOTIDE STR LOCI

A Thesis

Presented to

The Faculty of the Department of Biology

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Cynthia Renae Hall

December 2000

UMI Number: 1402510

Copyright 2000 by
Hall, Cynthia Renae

All rights reserved.

UMI[®]

UMI Microform 1402510

Copyright 2001 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2000

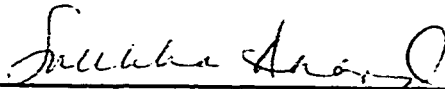
Cynthia Renae Hall

ALL RIGHTS RESERVED

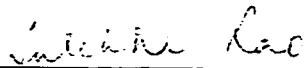
APPROVED FOR THE DEPARTMENT OF BIOLOGICAL SCIENCES

A handwritten signature in cursive script, appearing to read "David Matthes", written above a horizontal line.

Dr. David Matthes

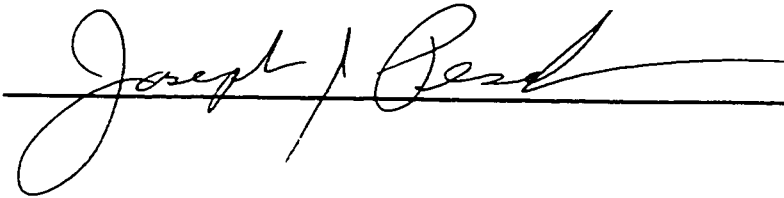
A handwritten signature in cursive script, appearing to read "Sulekha Anand", written above a horizontal line.

Dr. Sulekha Anand

A handwritten signature in cursive script, appearing to read "Sulekha Rao", written above a horizontal line.

Dr. Sulekha Rao, Applied Biosystems

APPROVED FOR THE UNIVERSITY

A large, stylized handwritten signature in cursive script, appearing to read "Joseph P. Bend", written above a horizontal line.

ABSTRACT

EFFECTS ON STUTTER PRODUCTION DURING AMPLIFICATION OF TETRANUCLEOTIDE STR LOCI

by Cynthia Renae Hall

Amplification of tetranucleotide short tandem repeat (STR) loci is routinely used for human identification purposes in forensic science to give highly discriminating results. Amplification of these STR loci commonly produces a minor PCR product termed “stutter”, which sizes one repeat unit shorter than its corresponding allele. This minor component peak is visible on the resulting electropherogram and may present a problem when interpreting body fluid mixtures typically encountered in forensic evidence samples. Locus size and flanking sequence, GC base content of core repeats, and primer position were examined in order to determine the effect of each on stutter production. The percent stutter produced increased as the size of the amplicon increased. This was caused by both increasing allele length and flanking sequence. Although some inhibition was encountered during amplification of DNA containing >50% GC, in general, percent stutter was also found to increase with GC content up to 50%, after which it decreased.

ACKNOWLEDGMENTS

I would like to thank my graduate committee: Dr. David Matthes, Dr. Sulekha Anand, and Dr. Sulekha Rao of Applied Biosystems. I would also like to express my gratitude and appreciation to Sean Walsh of Genometrix for his advice, contributions, and support on this project. I would like to thank Cassandra Musgrave-Nelson for her time and assistance with sample preparation and amplification. Laboratory facilities, equipment, and support were provided by the Santa Clara County Crime Laboratory, San Jose, CA and supplies and reagents were generously contributed by Applied Biosystems, Foster City, CA.

TABLE OF CONTENTS

List of Tables.....	viii
List of Figures.....	ix
Introduction.....	1
Forensic DNA Analysis.....	1
Short Tandem Repeats.....	4
Stutter.....	7
Materials and Methods.....	10
Human Subjects.....	11
DNA Samples and Extraction.....	11
DNA Amplification.....	13
Sample Electrophoresis and Data Collection.....	15
Experiment 1: Loci of the same repeat sequence.....	16
Data Analysis.....	18
Results.....	19
Experiment 2: Varying primer positions.....	22
Data Analysis.....	24
Results.....	24
Experiment 3: GC content of core repeat motif.....	27
Data Analysis.....	28
Results.....	28

Discussion.....	30
Conclusions.....	32
Appendix I: D5S818 Sequence.....	34
Appendix II: Human Subjects Approval.....	35
References.....	36

LIST OF TABLES

Table 1. Taq Gold Master Mix.....	14
Table 2. Stoffel fragment Master Mix.....	15
Table 3. Allele Information.....	17
Table 4. Primer Sequences.....	23
Table 5. Oligonucleotide Sequences.....	27

LIST OF FIGURES

Figure 1.	Slipped Strand Mismatching Model.....	8
Figure 2.	Electropherogram of Typical Stutter Peak.....	10
Figure 3.	Graph of Experiment 1: Main Effect of Allele with Taq Gold.....	20
Figure 4.	Graph of Experiment 1: Main Effect of Allele with Taq Stoffel.....	20
Figure 5.	Graph of Experiment 1: Main Effect of Locus with Taq Gold.....	21
Figure 6.	Graph of Experiment 1: Main Effect of Locus with Taq Stoffel.....	21
Figure 7.	Graph of Experiment 2: Interaction Between Primer Set and Allele with Taq Gold.....	25
Figure 8:	Graph of Experiment 2: Main Effect of Allele with Taq Stoffel.....	26
Figure 9:	Graph of Experiment 2: Main Effect of Primer Set with Taq Stoffel.....	26
Figure 10:	Graph of Experiment 3: Main Effect of GC Content with Taq Stoffel.....	29
Figure 11:	Graph of Experiment 3: Main Effect of GC Content with Taq Gold.....	29

Forensic DNA Analysis

The issue of human identification has always been an area of concern in the field of forensic science, and the analysis of biological evidence is no different. Biological fluids are key elements in the investigation of many types of crime. For example, in sexual assault cases, the assailant(s) may leave semen in the victim, on clothing, or at the scene (Reynolds *et al.*, 1991) while the victim may leave skin cells from saliva or vaginal fluid on the assailant. In homicide cases, the victim's blood may be found on the suspect's clothing or on a weapon found in the suspect's possession. The purpose of the serological examination is to determine, through genetic marker typing, whether the biological evidence left at a crime scene could have originated from a particular individual, either victim or suspect (Reynolds *et al.*, 1991). In the early days of conventional serology, this question was addressed by the use of antigen-antibody agglutination in the ABO blood group system and soluble polymorphic protein markers separated by electrophoresis. Some of these markers included red blood cell enzymes and serum proteins (Brewer, 1995) such as Phosphoglucosmutase (PGM₁), Esterase D (ESD), and Erythrocyte acid phosphotase (EAP or ACP₁). These markers are not very discriminating because they each only possess two to four different alleles. This does not allow the analyst to make any conclusive statements as to the identity of the individual depositing the evidence stain. Another inherent problem encountered with these markers is in the persistence of protein activity in a dried or liquid stain once it has left the body. Even under proper preservation (i.e. dried and frozen), these proteins can only be detected for a matter of a few weeks and sometimes up to a year. This does not allow for

testing of evidence associated with old crimes or those that either may not have been discovered right away or were subjected to harsh environmental factors. In addition, some of the markers used are limited in their body fluid detection. Of the many protein markers found in blood, only a small number of them (approximately ten) are present in other kinds of biological evidence such as semen and saliva (Reynolds *et al.*, 1991). Also, use of these markers does not allow for the separation of body fluid mixtures such as the vaginal fluid/semen mixtures found in sexual assault evidence.

Some of the problems encountered with conventional serological methods were overcome with the introduction of DNA analysis into the forensic arena. Forensic DNA analysis began with the use of restriction fragment length polymorphism (RFLP). This method is powerful in its ability to differentiate individuals and allows for differentially extracting DNA to separate spermatozoa from epithelial cells in mixed samples. However, RFLP also has certain limitations in its application to forensic science. The quantity and quality of DNA in a particular biological sample plays a very important role in determining how the sample should be analyzed (Whitaker *et al.*, 1995). Many samples encountered in forensic biology casework are either partially or severely degraded, limited in quantity, and several years old, making it almost impossible to analyze with RFLP. Restriction fragment length polymorphism requires a high quantity of non-degraded DNA, and since the nature of forensic samples rarely fits these criteria the application of this analysis is limited (Roy *et al.*, 1996 and Whitaker *et al.*, 1995). In order to address this problem, polymerase chain reaction (PCR) was introduced into the field of forensics. This methodology is successful with a limited amount of sample and

even degraded samples are successfully amplified, yielding some information as to the possible genetic identity of the individual. For example, a single hair found at a crime scene may be analyzed using PCR where it could not have been with RFLP.

There are several commercially available kits that have been used in the field of forensics for the analysis of various polymorphic loci, including AmpliType® HLA (Human Leukocyte Antigen) DQ- α , AmpliType® PolyMarker and D1S80 (Perkin Elmer). Both HLA DQ- α and PolyMarker (multiple loci co-amplified) detect nucleotide sequence polymorphisms by a reverse dot blot typing method. Although these two systems are effective at typing degraded samples, the discriminating power offered by them is low. HLA DQ α for example possesses only 8 different alleles, giving only 36 different possible genotypes. PolyMarker is somewhat more discriminating than DQ α with its ability to test multiple loci at once, however the two systems combined only give discrimination up to approximately 0.0003 (Whitaker *et al.*, 1995). This offers more information as to the individual identity than the conventional serological markers, but it is far from that offered by RFLP. D1S80 is a variable number tandem repeat (VNTR) locus with a 16 base pair repeat sequence that is amplified by PCR and electrophoresed on a polyacrylamide gel, followed by silver staining for allele detection. This system detects 27 different alleles, giving hundreds of possible genotypes. It is therefore much more discriminating than the reverse dot blot methods. However, due to the wide range of allele lengths associated with this locus, problems with amplification can occur. If a sample being tested has any of the higher molecular weight alleles, it is possible for amplification to fail due to lack of sufficient intact target molecules (Gill *et al.*, 1992 and

Walsh *et al.* 1992). In addition, there is an inverse relationship between amplification efficiency and allele length, which can cause preferential amplification (Demers *et al.*, 1995) and possibly lead to incorrectly calling a partially degraded heterozygote as a homozygote if the two alleles differ greatly in size (Whitaker *et al.*, 1995). Although all of these DNA typing systems have various benefits to the forensic community, they also have certain limitations. The current methodology being used is the amplification of short tandem repeats (STRs), which also has the same benefits associated with the other PCR based systems, but lacks most of the limiting factors.

Short Tandem Repeats

In the past several years, abundant research has described the presence of numerous stretches of highly polymorphic, repetitive DNA sequences scattered throughout the human genome (Kimpton *et al.*, 1993; Sprecher *et al.*, 1996; Fregeau and Fourney, 1993; Buel *et al.*, 1998). These consist of approximately one to seven bases arranged in a tandem array of widely varying lengths, which are known as microsatellites or STRs. Although these microsatellites are very abundant, their origin and functional significance (if any) is not currently understood (Levinson and Gutman, 1987). STRs are similar to the larger VNTR loci; however, they consist of smaller tandemly repeated sequences of one to seven base pairs, and allele lengths generally less than 350 base pairs. Unlike the VNTRs, STRs are found in regions other than telomeric regions and are also found in genic as well as extragenic areas (Kimpton *et al.*, 1993).

STR technology has been a significant advancement for forensics because of their increased amplification sensitivity, short analysis times, ease of interpretation (Sprecher *et al.*, 1996), and high abundance and polymorphism. Their small size allows for even the oldest, most degraded, and small quantity samples to be amplified without the preferential amplification and allelic dropout common to VNTR systems (Kimpton *et al.*, 1993). Due to the short lengths, alleles can be resolved to a single base pair difference, allowing for accurate allele designations without the use of the continuous allele distribution model necessary for VNTR analysis (Gill *et al.*, 1990 and Evett and Gill, 1991). The high abundance of STRs in the genome allows for several different loci to be co-amplified (multiplexed), resulting in more discriminating data for human identification and conservation of small samples. The use of fluorescently tagged primers allows, at present, up to nine loci, plus an X-Y homologous gender determination gene (Amelogenin) to be amplified in a single reaction tube. Loci with overlapping allele sizes can be multiplexed by strategic arrangement of primers with non-overlapping fluorescence maxima (Smith, 1995). This also permits co-electrophoresis of samples with an internal size standard, which minimizes lane-to-lane errors and allows for automated analysis, as well as automated allele calling when used in combination with well-characterized allelic ladders (Smith, 1995; Lins *et al.*, 1996; Fregeau and Fourney, 1993). The use of these primers also gives quantitative data on the relative intensities of the allele bands or peaks, which translates to the relative amounts of DNA present in a sample. This quantitative information aids in the interpretation of biological mixtures (Clayton *et al.*, 1998) commonly found in forensic samples. STRs are amenable to

detection through either gel-based electrophoresis and fluorescent detection using Perkin Elmer's ABI Prism™ 377 DNA Sequencer, or by capillary electrophoresis on the ABI Prism™ 310 Genetic Analyzer (Perkin Elmer).

Some of the most common, simple repetitive DNA sequences encountered consist of long polypyrimidine and poly-CA tracts, along with the tandem arrangements of closely related motifs (Levinson and Gutman, 1987) used in human identification. These sequences are found throughout the genomes of all eukaryotic organisms (Schlotterer and Tautz, 1992; Levinson and Gutman, 1987), and are thought to occur by a process known as slipped-strand mispairing. This mechanism involves local denaturation and displacement of the strands of a DNA duplex followed by mispairing of complementary bases at the site of an existing short tandem repeat. The simplest consequences of this mispairing, when followed by replication or repair, can lead to insertions or deletions of one or several of the short repeat units (Levinson and Gutman, 1987). There is increased opportunity for slippage to occur with each addition of a repeat sequence. It is for this reason that slippage during replication is thought to be the likely cause of the observed length polymorphism of simple sequence stretches between individuals within a population (Schlotterer and Tautz, 1992). This is a common phenomenon that is associated with several heritable human diseases such as Fragile X Syndrome and Myotonic Dystrophy. Both of these disorders are caused by expansion of a trinucleotide repeat array within the particular locus as the gene is passed to subsequent progeny, due to the formation of intermediate slipped strand DNA structures (Pearson *et al.*, 1998; Nolin *et al.*, 1999; Pearson and Sinden, 1996; and White *et al.*, 1999).

Stutter

The PCR amplification of short tandem repeat loci typically produces a minor product band one repeat unit shorter than the corresponding main allele band (Walsh *et al.*, 1996). These minor bands are commonly referred to as “stutter” or “shadow” bands, and are believed to occur through the same slippage mechanism that is responsible for repetitive polymorphisms and the manifestation of various genetic disorders. In the case of PCR amplification, slippage can occur during the process of chain elongation by *Taq* DNA polymerase. If there is a pause during elongation, the DNA polymerase may become detached from the DNA. Strand separation can then occur followed by looping out of the template strand and ‘jumping’ of an integral number of repeat units (Murray *et al.*, 1993) (Figure 1). This process creates a product which is shorter than the main allele by multiples of the repeat unit; however, it is also possible (at lower levels) for the extending strand to loop out and produce shadow bands which are longer than the main allele. It has been shown that within a given STR locus, the proportion of stutter produced during PCR is highest for alleles that contain long stretches of uninterrupted core repeat sequence (the most common repeat motif within the variable expansion region) and generally increases with allele length (Walsh *et al.*, 1996). Magnesium concentration and temperature may also play a role in the amount of stutter produced. In addition, it has been suggested that slippage is potentially dependent on the AT-content of the sequences involved (Schlotterer and Tautz, 1992) and that a thermostable DNA polymerase with greater processivity than *Taq* polymerase may reduce the opportunity

for the formation of slipped-DNA during PCR, and thus reduce stutter bands (Murray *et al.*, 1993; Schlotterer and Tautz, 1992; Walsh, *et al.* 1996).

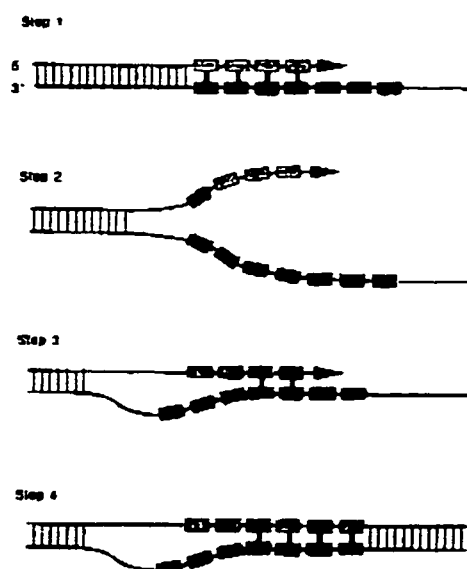


Figure 1. Slipped strand mispairing model. Diagrammed in this figure is a proposed mechanism to explain the formation of stutter bands one repeat unit shorter than the main allele band. The template strand (main allele) in this figure contains seven repeat units as represented by the dark gray shaded boxes. In Step 1, the DNA polymerase has extended through four repeat units, represented by the light gray boxes. If /when the polymerase falls off the extending strand, the template and extending strand can breath apart as shown in Step 2. When the two strands re-anneal in Step 3 the template strand has looped out and the extending strand aligns out-of-register by one repeat unit. Thus the polymerase can only add an additional two repeat units, instead of the correct three repeat units. Step 4 demonstrates that at the completion of synthesis the newly extended strand contains only six repeat units, while the template strand has seven repeat units. (Figure taken from Walsh *et al.*, 1996)

Although the stutter phenomenon is present in di-, tri-, and tetranucleotide repeats, it is more prevalent with the dinucleotide sequence motifs. A disadvantage of the dinucleotide repeat polymorphism is that each allele consists of not a single fragment, but a series of multiple shadow bands (Litt *et al.*, 1993). The most common stutter band is

typically two bases shorter than the main allele; however, additional bands are also present four and six bases shorter than the main allele (Walsh *et al.*, 1996). This multi – banding pattern poses a problem in the interpretation of genotypes commonly encountered with forensic specimens and also in samples from a heterozygote whose alleles are very close in size. Tetranucleotide repeats also produce stutter bands when amplified by PCR, however the minor product usually consists of only a single band located four base pairs shorter than the main allele (Figure 2), and to a lesser degree than that of the dinucleotides, making genotype interpretation easier. Various tetranucleotide loci have been chosen for forensic DNA analysis due to the lower amount of stutter and easier interpretation of shadow banding patterns, as opposed to the di- and trinucleotides. However, there can still be problems in distinguishing between stutter and true minor alleles that may be present in the mixed samples common in forensic specimens, due to the fact that stutter bands are the same size as true alleles. Quantitative information obtained in the analysis of the STRs can be used to aid in this distinction given that the percent stutter (stutter peak amplitude relative to allele peak amplitude) observed at any given locus should be less than 15 – 18% for a single source sample (Perkin Elmer AmpF/STR Profiler™ and Profiler Plus™ User's Manuals). Despite this general guideline, it is still desirable to further reduce and if possible eliminate stutter in the analysis of STRs for human identification purposes.

Factors affecting the production of stutter during amplification of STRs were identified in this project by examining the effects of primer and flanking DNA sequences,

primer position in relation to core repeat sequence, and varying GC content within a given tetranucleotide repeat motif.

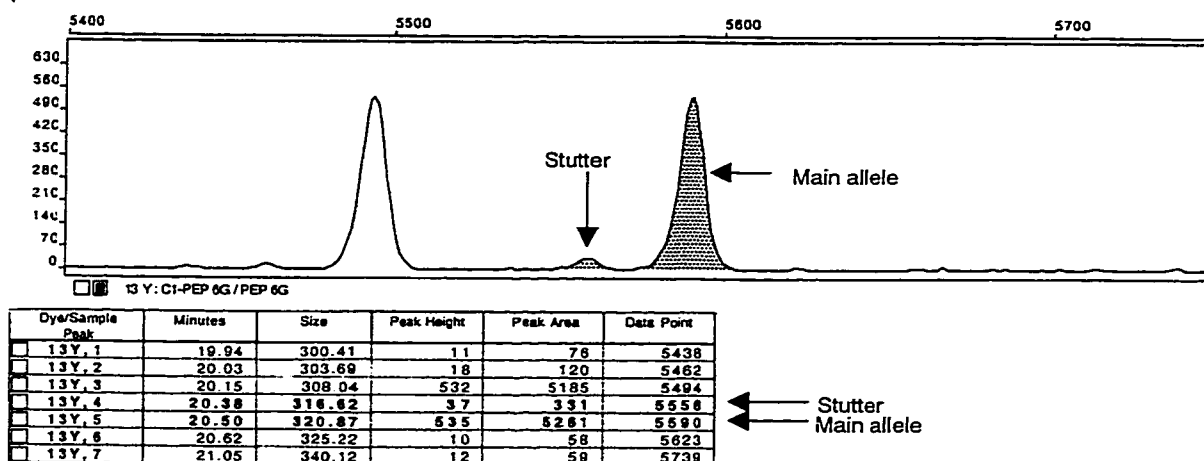


Figure 2. Electropherogram depicting a typical stutter peak observed following amplification of a tetranucleotide STR locus. This sample was amplified at the CSF1PO locus using *Taq* polymerase gold. The shaded peaks represent a main allele containing 13 repeat units (320.87 bp) and its corresponding stutter peak (316.62 bp) located one repeat unit (4 bp) shorter than the main allele.

MATERIALS AND METHODS

The approach for these experiments was to first determine the effect of primer and flanking DNA sequences on the amount of stutter produced. This was done by examining two loci (D5S818 and CSF1PO) and an oligo template with the same core repeat number and sequence motif. Next, in order to test primer position in relation to stutter production, the D5S818 locus was used to design two additional primer sets which would give one amplicon which was shorter and one longer than that normally produced with the current primers. Finally, additional oligos were designed which varied from the

first oligo template in the GC content present in the repeat sequence, followed by comparison of the amount of stutter amplified for each.

Human Subjects

Chromosomal DNA samples used in this study consisted of biological specimens previously obtained from individuals for the purpose of DNA method validation and population data studies conducted at the Santa Clara County Crime Laboratory. DNA from each of the specimens was previously extracted and quantitated as part of the original studies and were chosen for use based on previously determined genotypes. Laboratory officials granted permission for use of these samples for the purpose of conducting research, so long as subject confidentiality was maintained. In order to ensure that personal information was kept confidential, specimen numbers originally assigned as part of the Crime Laboratory's endeavors were used in this research for identification purposes. All of the identifying information associated with each sample is maintained in locked filing cabinets located within the laboratory. Only members of the laboratory's managerial staff have access to these files. Due to the nature of the samples being used, there were no other risks associated to the individuals involved.

DNA Samples and Extraction

All of the genomic samples used consisted of DNA previously extracted according to the Santa Clara County Crime Laboratory's procedure manual, using a Phenol-chloroform organic extraction with a Centricon YM-100 (Millipore, Bedford,

MA) concentration, or Chelex® 100 Resin (Bio-Rad Laboratories, Hercules, CA). These extractions were performed as part of the laboratory's population database study and/or validation of the AmpF/STR Profiler Plus™ and COfiler™ Amplification Kits (Perkin Elmer). Briefly, the organic extraction procedure was conducted as follows: samples were added to 0.5mL digest buffer (10mM Tris-HCL, 10m EDTA, 50mM NaCL, 2% SDS, pH 7.5) and 15 µL of 10 mg/mL Proteinase K, incubated at 56°C overnight, followed by addition of buffered phenol-chloroform solution x2 and Centricon-100 concentration (centrifuge in fixed angle rotor at 2800 RPM for 15 min x3, and retentate collection at 1800 RPM for 2 min). The samples extracted using the Chelex procedure were incubated in 1 mL sterile distilled water at room temperature for 15 to 30 min, centrifuged for 3 min at 10,000 to 15,000 x g. The supernatant was discarded and 5% Chelex solution (Chelex resin beads suspended in sterile water) was added to a final volume of 200 µL. This mixture was incubated at 56°C for 15 to 30 min, vortexed briefly and boiled in a water bath for 8 min, followed by centrifugation for 3 min.

Following extraction, samples were quantitated using the QuantiBlot® Human DNA Quantitation Kit (Perkin Elmer), according to the manufacturer's recommendations. This procedure consists of the hybridization of a biotinylated oligonucleotide probe to DNA samples immobilized on a nylon membrane. The probe is complementary to a primate-specific alpha satellite DNA sequence at the D17Z1 locus. Detection is performed by the binding of enzyme:conjugate:HRP-SA (horseradish peroxidase-streptavidin) to the biotin moiety of the probe, which results in a blue-colored precipitate formation through oxidation of chromogen:TMB, catalyzed by horseradish

peroxidase. Genotypes were then determined at thirteen loci plus Amelogenin using ABI Prism Genotyper® Software. Samples used in this research project were chosen based on these previously determined genotypes.

DNA Amplification

In order to increase the amount of stutter obtained for each sample and possibly more effectively interpret any differences that may be occurring, all of the studies conducted were run in duplicate. All of the samples were first amplified, using AmpliTaq® Gold DNA Polymerase (Applied Biosystems), according to the Santa Clara County Crime Laboratory's protocol and the AmpF/STR Profiler Plus™ and COfiler™ User's Manuals. Next, the same samples were amplified using Stoffel fragment of AmpliTaq® DNA polymerase (Applied Biosystems), as recommended by the manufacturer's product insert. Stoffel fragment is a truncated *Taq* polymerase in which 289 amino acids have been deleted and which has been shown to cause a three-fold increase in stutter (Walsh *et al.*, 1996). This step was performed in order to possibly make any small differences more readily apparent. All of these amplification reactions were performed in GeneAmp® Thin-Walled Reaction Tubes in a Perkin Elmer DNA Thermal Cycler 480. The cycling parameters for samples amplified with Taq Gold were set at 95°C for an initial 11 min. incubation, 94°C for 1 min. (denature), 50°C for 1 min. (anneal), and 72°C for 1 min. (extend), followed by a final extension step, after 28 cycles, for 45 min. at 60°C. The parameters for the Stoffel reactions were the same as those for Taq Gold except that the initial 11 min activation step at 95°C was eliminated.

Genomic samples used for this study were amplified using a target range of 2.5 to 3.0 nanograms (ng) of DNA in 50 μ L reaction volumes. Oligonucleotides were amplified with a target of approximately 300 to 500 molecules per 50 μ l reaction. PCR reaction mix consisting of $MgCl_2$, deoxynucleoside triphosphates (dATP, dCTP, dGTP, dTTP), 160 μ g/ml bovine serum albumin (BSA) and 0.05% sodium azide (NaN_3) in buffer and salt (AmpF/STR Profiler Plus™ User's Manual), and 15 pmol fluorescently tagged locus specific primers (see Table 2 for primer sequences) were used for the Taq Gold reactions (Table 1).

Table 1. Taq Gold Master Mix^a

<u>Component</u>	<u>Volume</u>
reaction mix	21 μ l x (X)
AmpliTaq Gold	1 μ l x (X)
forward primer*	6 μ l x (X)
reverse primer	6 μ l x (X)

X = number of samples to be amplified

* = fluorescently labeled primer

- a. 30 μ l master mix and 20 μ l DNA were added to give a final volume of 50 μ l with one drop of mineral oil on top

A test set of amplifications were performed using the AmpliTaq® DNA polymerase Stoffel Fragment (Applied Biosystems) according to the manufacturer's recommendations of a 100 μ l reaction volume and with a final volume of 50 μ l. Both sets of amplifications were successful with equal peak intensities; therefore the reduced volumes of 50 μ l were used for this study according to the master mix listed in Table 2.

Table 2. Stoffel Fragment Master Mix^a

<u>Component</u>	<u>Volume</u>	<u>Final Concentration</u>
10X Stoffel Buffer	10 μ l (1/2X)	5 mM Tris-HCl, pH 8.3, 5 mM KCl
dNTP mix	16 μ l (1/2X)	2.5 mM (.625 mM each dNTP)
forward primer*	6 μ l (1/2X)	7.5 pmol
reverse primer	6 μ l (1/2X)	7.5 pmol
Stoffel fragment	.5 μ l (1/2X)	2.5 Units
25 mM MgCl ₂	20 μ l (1/2X)	2.5 mM
TE buffer	21.5 μ l (1/2X)	

X = number of samples to be amplified

* = fluorescently labeled primer

a. 40 μ l master mix and 10 μ l DNA were added to give a final volume of 50 μ l with one drop of mineral oil on top

Sample Electrophoresis and Data Collection

Capillary electrophoresis was performed on the ABI Prism 310 Genetic Analyzer. The Genetic Analyzer protocol was followed as outlined in the User's Manuals: 1.5 μ L amplified product and 1 μ L of GeneScan-500 [ROX] Internal Lane Size Standard was added to 24 μ L deionized formamide (purchased from Amresco) and denatured at 95°C for 3 min, followed by snap cooling for 3 min in an ice bath. The PCR products were injected for 5 sec at 60°C and electrophoresed at 15 kV in Performance Optimized Polymer 4 (POP4), consisting of linear dimethylacrylamide, 8 M urea, 5% 2-pyrrolidinone, and 1 mM EDTA (Lazaruk *et al.*, 1998). As samples move past a laser detection window, their laser-induced fluorescence was detected on a CCD camera. Data were then collected using the ABI Prism 310 Collection software with the run module GS STR POP4 (1 mL) F, and results analyzed using the 310 GeneScan software application. The electropherograms produced and the corresponding data tables were then used to identify the allele and stutter of interest and calculate the peak height ratios (% stutter) for

each. The samples were analyzed in GeneScan using light smoothing and a peak height threshold cutoff of 150 RFU for the preliminary allele identification. The peak height threshold was then lowered to 10 RFU in order to obtain peak heights for low stutter peaks. If the stutter peak was indistinguishable from the baseline background, the amount of stutter for that particular sample was scored as zero.

Experiment 1: Loci of the same repeat sequence

The first question addressed in this project was whether the stutter artifact produced during PCR amplification of STRs is influenced by locus and allele length. This was explored by amplifying samples of previously determined genotypes at the D5S818 and CSF1PO loci. The amplifications were performed on samples containing alleles of the same number of consecutive repeats in order to examine the amount of stutter produced at different loci containing core repeats of the same nucleotide sequence. The sequence motifs for the loci of interest are illustrated in Table 3. Both CSF1PO and D5S818 contain a simple core repeat motif of AGAT, however D5S818 produces an amplicon of approximately 130 – 171 base pairs, and CSF1PO produces a longer amplicon of approximately 281 – 320 base pairs, depending on the number of repeats per allele.

DNA extracted from 30 individuals was obtained based on each of their previously determined genotypes at the CSF1PO and D5S818 loci and the amount of DNA present in the extract. Fifteen of the individuals were chosen for their CSF1PO types and fifteen for D5S818. In order to efficiently determine the amount of stutter

produced for each allele, it was important that the sample chosen met certain criteria. Each individual was heterozygous at the particular locus and his/her alleles were at least two repeat units (8 base pairs) apart. This is because the amount of stutter observed may be higher for homozygote samples due to the increased peak height in relation to heterozygote samples. Additionally, stutter may be masked by alleles less than eight base pairs apart, as the allele and the stutter peak reside in the same location.

The fifteen subjects for each locus were grouped according to their previously determined genotypes. Five samples contained a seven allele (seven repeats), five possessed a ten allele (ten repeats), and the other five possessed a 13 allele (13 repeats). Each sample was amplified using both AmpliTaq Gold and again using AmpliTaq Stoffel Fragment, for a total of five amplifications per sample for each enzyme. For each subject, the peak height ratios (percent stutter) obtained for each of the five amplifications were averaged together, producing one data point per person. One of the samples with a seven allele at D5S818 did not contain enough DNA; therefore only four samples were used for the Stoffel amplifications.

Table 3. Allele Information^a

Locus	Chromosomal Location	Core Repeat Sequence Motif	Allele Size Range
D5S818	5q21-31	(AGAT) _n	135 -171 bp 7 - 16 repeats
CSF1PO	5q33.3-34	(AGAT) _n	281 - 317 bp 6 - 15 repeats

a. Information obtained from the AmpF/STR Profiler and Profiler Plus User's Manuals (Perkin Elmer)

In addition to the chromosomal DNA used in this experiment, a test oligonucleotide was manufactured for use as a template in order to determine whether

stutter amplified from an oligo target is comparable to that of chromosomal target DNA. This step was an important one to establish, as several manufactured templates were used exclusively in the third part of this research project. The oligo contained ten repeat units of the same sequence as D5S818 and CSF1PO (AGAT), along with truncated flanking sequence and primers for D5S818. This test oligo was amplified in replicate (five amplifications, giving five data points) and the amount of stutter compared to that obtained for the individuals with ten repeat alleles at both D5S818 and CSF1PO. The averaged data was used for the genomic samples, producing five data points for each locus.

Data Analysis

The percent stutter obtained for each allele was calculated by determining the peak heights (measured in relative fluorescence units) for both the main allele and its corresponding stutter product. The peak height ratios (%) were determined by dividing the stutter peak by the allele peak. Mean peak height ratios (mean of the five data points) were determined for each individual within each allele and results analyzed in SYSTAT (ver. 9.0) using analysis of variance (ANOVA). Locus and allele were used as the independent variables and percent stutter (mean peak height ratio) was the dependent variable. In order to reduce the chance of making a Type I error and mistakenly concluding that there were effects on stutter, alpha was set at 0.025.

The five peak height ratios calculated for the test oligo were compared to the mean peak height ratios calculated for each of the five individuals possessing an allele

length of ten at D5S818 and CSF1PO. These peak height ratios were also analyzed using ANOVA in SYSTAT (ver. 9.0) with alpha set at 0.025. Locus/Oligo was used as the independent variable and percent stutter was the dependent variable.

Results

All of the stutter peaks produced were located in a position four base pairs shorter than the corresponding main allele peak, as expected for tetranucleotide repeat loci. A few samples did not produce any stutter for a particular amplification, or the peak was too low to distinguish it from baseline noise. These samples were most prevalent when analyzing those containing a seven allele at D5S818 when amplified with Taq Gold. In general, the amount of stutter produced for samples amplified using Taq Gold was lower than that produced when amplified with Taq Stoffel, as had been shown in previous studies (Walsh *et al.*, 1996).

Regardless of which enzyme was used for amplification, there was no interaction between allele length and locus (Gold: $F(2, 24) = 3.491$, $p = 0.047$, Stoffel: $F(2, 23) = 1.426$, $p = 0.261$). Allele length had the same effect for both loci. ANOVA with allele and locus as the independent variables also indicated a main effect of allele on percent stutter when amplified with Taq Gold ($F(2, 24) = 62.551$, $p < 0.001$) and with Taq Stoffel ($F(2, 23) = 635.745$, $p < 0.001$). The amount of stutter produced during amplification with either enzyme increased with allele length (Figures 3 and 4). This was the same trend previously reported by Walsh, et al., (1996).

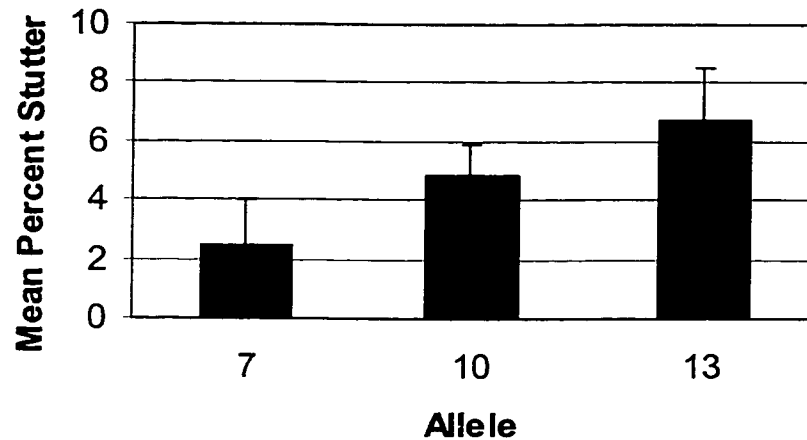


Figure 3. Effect of allele on percent stutter when amplified with AmpliTaq Gold ($p < 0.001$). Each allele length represents mean percent stutter results for both CSF1PO and D5S818 combined. The error bars equal + one standard deviation from the mean.

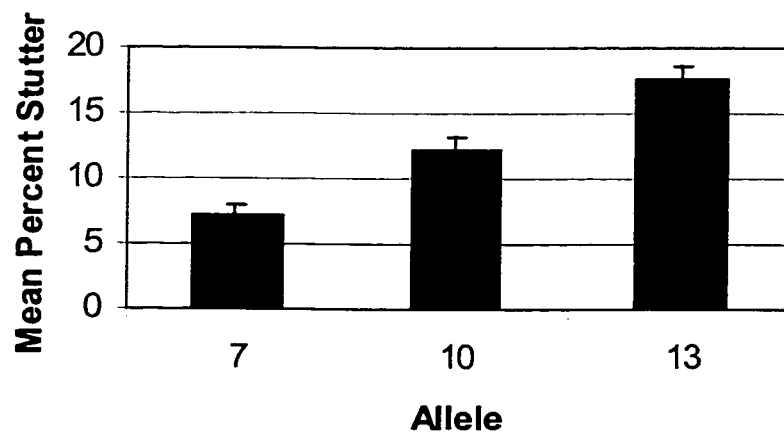


Figure 4. Effect of allele on percent stutter when amplified with AmpliTaq Stoffel ($p < 0.001$). Each allele length represents mean percent stutter results for both CSF1PO and D5S818 combined. The error bars equal + one standard deviation from the mean.

In addition to the effect of allele length on stutter, ANOVA also indicated a main effect of locus on the amount of stutter produced when amplified with Taq Gold ($F(1, 24) = 57.101, p < 0.001$) and with Taq Stoffel ($F(1, 23) = 25.585, p < 0.001$). CSF1PO

generally showed an increase in the amount of stutter observed over that of D5S818 (Figures 5 and 6). This increase in stutter for CSF1PO occurred for samples amplified with both Taq Gold and Taq Stoffel.

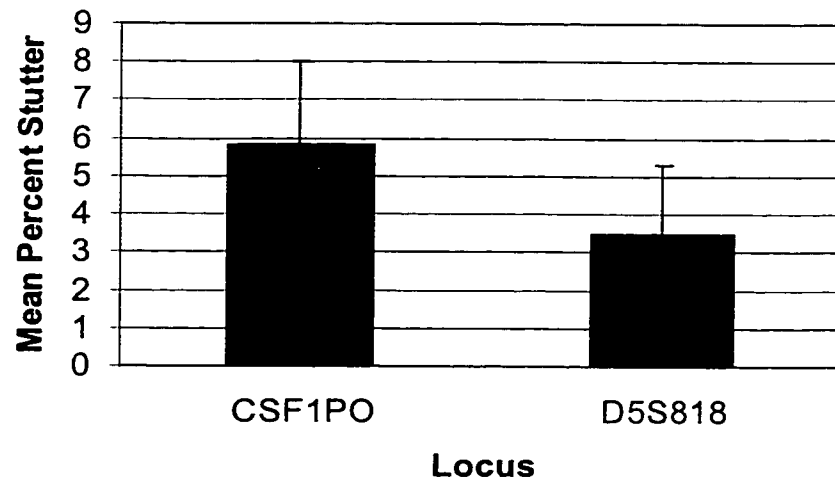


Figure 5. Effect of locus on percent stutter when amplified with AmpliTaq Gold ($p < 0.001$). Each locus represents mean percent stutter results for all three allele lengths combined. The error bars equal + one standard deviation from the mean.

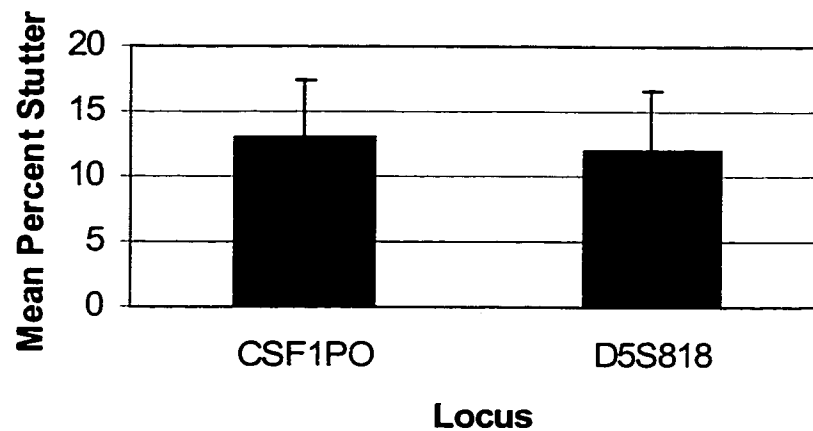


Figure 6. Effect of locus on percent stutter when amplified with AmpliTaq Stoffel ($p < 0.001$). Each locus represents mean percent stutter results for all three allele lengths combined. The error bars equal + one standard deviation from the mean.

Additionally, a test oligonucleotide containing ten repeats of the sequence AGAT and using D5S818 primer and truncated flanking sequence was amplified five times with both Taq Gold and Taq Stoffel. The percent stutter calculated for each of these five amplifications was compared to the mean percent stutter obtained for each of the individuals with an allele of 10 at D5S818 and CSF1PO. ANOVA indicated that there was no statistically significant difference between the genomic DNA and the oligonucleotide using either Taq Gold ($F(2, 12) = 1.248, p = 0.322$) or Taq Stoffel ($F(2, 12) = 4.860, p = 0.028$). This indicates that a manufactured oligo produces comparable amounts of stutter as chromosomal DNA samples and was an acceptable model for use in the third experiment of this research.

Experiment 2: Varying primer positions

The second question addressed in this study was whether primer position in relation to the core repeat sequence during amplification has an effect on the amount of stutter produced. This issue was addressed by examining the sequence information for the D5S818 locus (Genbank accession #: G08446) and its corresponding primers (primer set 1). Using this sequence information, two additional sets of primers located at varying distances from the common sequence motif (Table 4) were designed with Applied Biosystems' Primer Express software. One of these sets of primers was located as close to the core motif as possible to produce an amplicon shorter than that normally obtained (primer set 2). The other was positioned at a distance further away than the original primers to amplify a sequence as long as feasible (primer set 3). Melting temperature

(T_m) and length of the primers (set 2 and 3) were maintained as close as possible to that of primer set 1. The reverse primer for set 3 largely overlapped that for the original primer set. The possibilities for positioning this particular primer at a large distance from the repeat motif was limited due to the lack of known sequence information in this area of the locus (Appendix I).

DNA from the same set of individuals chosen for all three allele lengths at D5S818 in experiment one was amplified using these two new sets of primers and compared to the results obtained for the original primer set. Five replicate amplifications were performed with each of the three primer sets for each individual using both Taq Gold and Taq Stoffel. Peak height ratios for each of the five replicate amplifications was averaged for each subject, producing one data point per person for each primer set. One of the samples containing a seven allele did not have a sufficient amount of DNA present, so it was amplified only three times with each of the two new primer sets using Taq Gold. This sample was not amplified at all for any of the primer sets using Taq Stoffel.

Table 4. Primer Sequences^a

Locus	Primer Sequences	Source
D5S818 (primer set 1)	gggtgatttcctcttgggt tgattccaatcatagccaca	MarkerSearch Accession # 512
CSF1PO	aacctgagctcgccaaggactagc ttccacacaccactggccatcttc	Huang <i>et al.</i> , 1995
D5S818 (primer set 2)	cctctttggtatccttatgtaatt ctgtatccttattataacctctatct	Designed using Primer Express
D5S818 (primer set 3)	gggtgccagataatctgtactaata tttagccaagtgtattccaatcata	Designed using Primer Express

a. Primers manufactured and supplied by Applied Biosystems (Foster City, CA)

Data Analysis

The percent stutter was calculated as in experiment one. The mean peak height ratios were determined and the results analyzed using repeated measures ANOVA in SYSTAT (ver. 9.0), with primer set as a repeated measures variable, allele as a between-subjects variable, and percent stutter as the dependent variable. Alpha was again set at 0.025 to reduce a Type I error and mistakenly concluding that there was an effect on stutter.

Results

The original primer set (primer set 1) produced amplicons ranging from approximately 152 – 176 base pairs for alleles containing 7 – 13 repeat units. Primer set two, positioned as close to the core repeat motif as possible produced amplicons ranging from approximately 91 – 127 base pairs and primer set three (as far away as possible) gave a range of 209 – 233 base pairs.

Repeated measures ANOVA with primer set (repeated measure) and allele as independent variables indicated an interaction between primer set and allele ($F(4, 24) = 6.247$, $p = 0.001$) and main effects of primer ($F(2, 24) = 40.194$, $p < 0.001$) and allele ($F(2, 12) = 75.044$, $p < 0.001$) when amplified with Taq Gold. The same general trend of increasing stutter with increasing allele length was observed for each primer set. Primer set affected stutter differently depending on allele length as shown in Figure 7. Primer set 3 (longest amplicon) produced the highest amount of stutter for all three allele lengths, while primer sets 1 and 2 (shortest amplicon) varied slightly in stutter production

by allele. With the exception of alleles of length 13 however, the amount of stutter produced generally showed an increase as amplicon size increased.

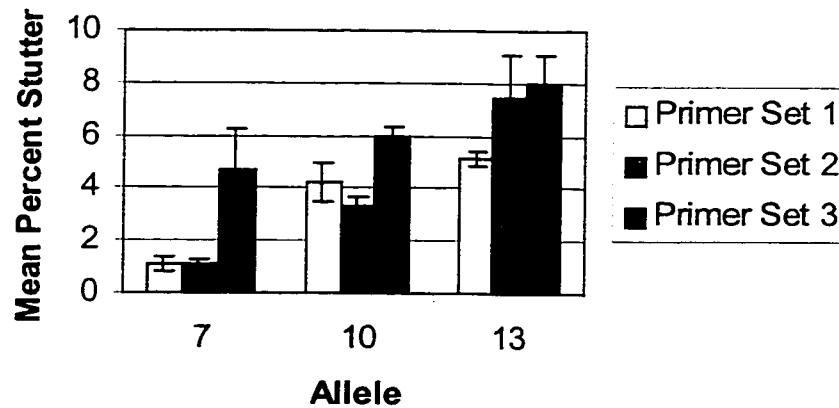


Figure 7. Interaction between primer set and allele when amplified using AmpliTaq Gold ($p = 0.001$). Error bars equal \pm one standard deviation from the mean.

When amplified with Taq Stoffel, repeated measures ANOVA with primer set (repeated measure) and allele as independent variables indicated main effects of primer ($F(2, 22) = 10.996$, $p < 0.001$) and allele ($F(2, 11) = 1718.259$, $p < 0.001$) and no statistically significant interaction between primer set and allele ($F(4, 22) = 2.311$, $p = 0.098$). The same general trend shown in previous studies was also observed with these sets of amplifications. The amount of stutter produced during amplification increased with larger allele sizes (Figure 8). All three primer sets had the same general effect on stutter regardless of allele length (Figure 9). The original set of primers (primer set 1) exhibited the lowest amount of stutter out of the three sets. However, the shortest

amplicons (primer set 2) produced a slight increase in stutter over primer set 3 (longest amplicon).

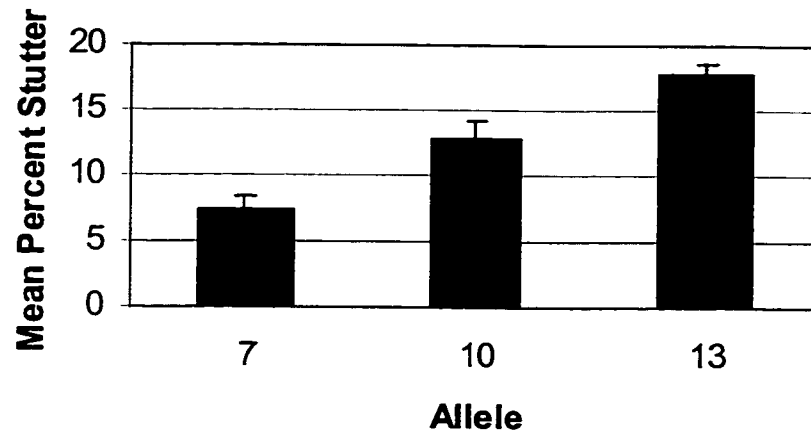


Figure 8. Effect of allele on percent stutter when amplifying various primers using AmpliTaq Stoffel Fragment ($p < 0.001$). Each allele represents mean percent stutter for all three primer sets combined. The error bars equal + one standard deviation from the mean.

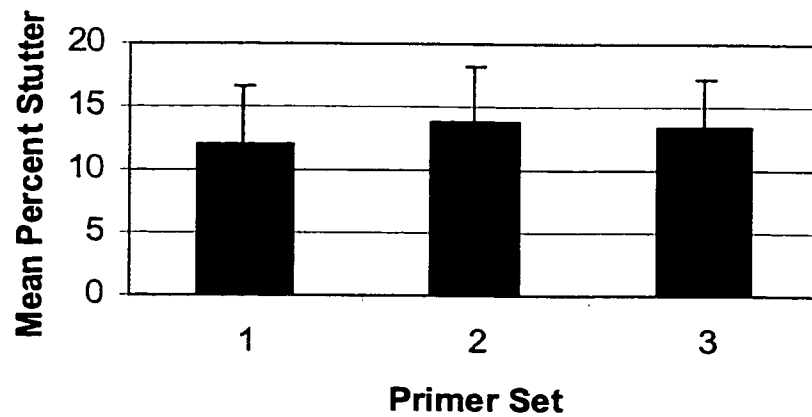


Figure 9. Effect of primer set on percent stutter when amplified with AmpliTaq Stoffel Fragment ($p < 0.001$). Each primer set represents mean percent stutter for all three allele lengths combined. Error bars equal + one standard deviation from the mean.

Experiment 3: GC content of core repeat motif

The final question addressed in this research was the effect of core repeat composition and content on stutter. Using the primers and flanking sequence for the test oligo described in experiment one, five additional oligonucleotides were manufactured to contain different tetranucleotide repeat sequences of varying GC content. All of these templates contained ten tetranucleotide repeat units; however, the GC content ranged from 0% (R2) to 100% (R6) in increments of 25% (See Table 5 for sequence). The six oligonucleotides, including the test oligo (R1) from experiment one were amplified five times each using both Taq Gold and Stoffel, producing five data points per oligo.

Table 5. Oligonucleotide Sequences^a

Oligonucleotide	Sequence
Test oligo (25%GC) R1	GGGTGATTTTCCTCTTTGGTatccttatgtatatttgaa (GATA) ₁₀ ggtagatagatggtgtaaacTGTGGCTATGATT GGAATCA
0% GC R2	GGGTGATTTTCCTCTTTGGTatccttatgtatatttgaa (AATT) ₁₀ ggtagatagatggtgtaaacTGTGGCTATGATT GGAATCA
25% GC R3	GGGTGATTTTCCTCTTTGGTatccttatgtatatttgaa (AAGT) ₁₀ ggtagatagatggtgtaaacTGTGGCTATGATT GGAATCA
50% GC R4	GGGTGATTTTCCTCTTTGGTatccttatgtatatttgaa (AAGG) ₁₀ ggtagatagatggtgtaaacTGTGGCTATGATT GGAATCA
75% GC R5	GGGTGATTTTCCTCTTTGGTatccttatgtatatttgaa (AGGC) ₁₀ ggtagatagatggtgtaaacTGTGGCTATGATT GGAATCA
100% GC R6	GGGTGATTTTCCTCTTTGGTatccttatgtatatttgaa (GGCC) ₁₀ ggtagatagatggtgtaaacTGTGGCTATGATT GGAATCA

a. Oligonucleotides manufactured and provided by Applied Biosystems

Data Analysis

The percent stutter was calculated as previously described for each of the five amplifications for each oligo. The test oligo (R1) and oligo 3 both contained 25% GC so results were combined into one category for data analysis. The results were analyzed using ANOVA in SYSTAT (ver. 9.0), with oligo (GC content) as the independent variable and percent stutter as the dependent variable. In order to reduce the chance of making a type I error, alpha was again set at 0.025.

Results

Some difficulty was met in amplifying oligos R5 and R6 (75% and 100% GC respectively). The product obtained for both of these oligos was either very weak, or nonexistent for both Taq Gold and Taq Stoffel. For each of the five amplifications, results were obtained for only three of the R5 stoffel samples, two of the R6 stoffel, and none of the R5 gold samples. All of the R6 gold samples were very weak and the percent stutter was scored as zero because it was indistinguishable from the baseline noise. Several attempts were made to amplify these samples, varying the concentration of template DNA and electrophoresis parameters, however the results did not differ from those being reported here.

ANOVA with oligo as the independent variable indicated a main effect of oligo when amplified with Taq Stoffel ($F(4, 20) = 76.844$, $p < 0.001$) and with Taq Gold ($F(3, 21) = 18.901$, $p < 0.001$). Stutter increased with GC content up to 50%, and then sharply decreased at 75% and 100% (Figures 10 and 11). The amount of stutter produced for

oligos R5 and R6 using Stoffel fragment is very similar, with R6 being slightly higher, again showing a general increase with increase in GC. No interpretation can be made as to the R5/R6 relationship using Taq Gold, as no results could be obtained for R5.

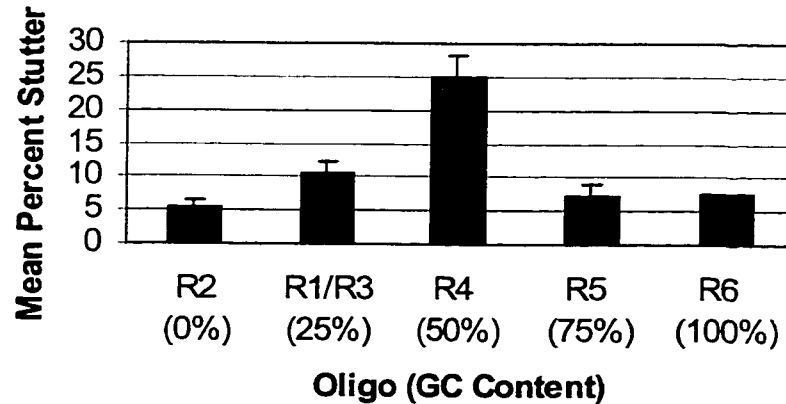


Figure 10. Effect of core GC content on percent stutter when amplified with AmpliTaq Stoffel Fragment ($p < 0.001$). Error bars equal + one standard deviation from the mean. R5 contained an $N = 3$, and R6 contained $N = 2$.

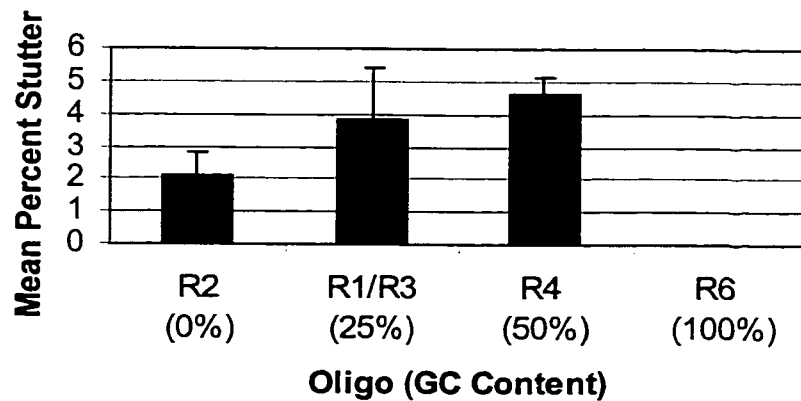


Figure 11. Effect of core GC content on percent stutter when amplified with AmpliTaq Gold ($p < 0.001$). Error bars equal + one standard deviation from the mean. No results were obtained for oligo R5.

DISCUSSION

Overall, the results of this research are consistent with previous results reported by Walsh *et al.*, (1996). The amount of stutter produced for a particular locus increases as allele length increases. Also, stutter is generally higher for the locus CSF1PO when compared to that of D5S818. This may be due to the same phenomenon that is occurring with allele length. CSF1PO typically produces DNA segments greater than 100 base pairs longer than those produced for D5S818. Following the proposed mechanism of slipped-strand mispairing, these longer allele lengths may present more opportunity for the enzyme to become dissociated from the template due to enzyme processivity. This disassociation creates a pause in the elongation step, which can allow the DNA to loop out in the repeat motif before the enzyme reestablishes itself to finish the nucleotide additions. This difference in the amount of stutter produced between loci also suggests that the length of the flanking DNA sequence may influence stutter production.

Although primer set (primer position) had an effect on the amount of stutter produced, there was a difference observed between samples amplified with AmpliTaq Gold and AmpliTaq Stoffel Fragment. The general trend observed for those using Taq Gold was that the amount of stutter produced increased as the size of the amplicon produced increased. In other words, the further the two primers are positioned away from the core repeat motif, the more stutter is produced. Assuming the mechanism of slipped-strand mispairing, this again may be due to more opportunity for the enzyme to “fall off” of the template, allowing the two strands to breath apart and form a loop structure in the repeat region.

One exception to this trend was observed for alleles containing 13 repeats in the Taq Gold samples. For this allele length, two sets of amplifications with primer set 2 produced weakly amplified product in comparison to the other three sets. These weak samples created an artificially high amount of stutter when analyzed, due to a much higher background noise level. These higher peak height ratios also caused an increase in the overall mean percent stutter observed for this allele and primer set, making it appear higher than it should have been. This artificially high mean may explain the difference in the results for this particular allele. The other difference observed in this experiment was in the samples amplified with AmpliTaq Stoffel. These samples showed an interesting result. Stutter produced was lowest for the original set of primers (set 1, mid length amplicon). The next highest amount was for the longest set of amplicons (set 3) and was highest for the shortest amplicons (primer set 2). Although samples amplified with Stoffel fragment show a three-fold increase in the amount of stutter produced (Walsh *et al.*, 1996), it has not been shown if this increase is amplified proportionally. This possible lack of proportion may explain the difference observed with Stoffel in this study.

Results of this research also showed that the GC content of the core repeat motif effects the amount of stutter produced. Although the phenomenon appears to be an increase in stutter with increased GC content, the difficulty in obtaining product for the two highest GC content oligos makes interpretation of the trend difficult beyond 50% GC. It has been shown in previous studies that samples with higher GC content inhibits amplification (McDowell *et al.*, 1998; Mytelka and Chamberlin, 1996; Sarkar *et al.*, 1990). These samples containing GC-rich domains give rise to stable secondary

structures and higher melting temperatures which could serve as a pause or termination site for the enzyme (McDowell *et al.*, 1998), inhibiting its ability to extend the primer. It was also proposed that these GC-rich regions might slow the polymerase to the point that elongation pauses due to difficulty in incorporating nucleotides (Mytelka and Chamberlin, 1996). This problem could explain the lack of amplified product obtained for oligos R5 (75% GC) and R6 (100% GC). Several methods have been proposed to overcome this enzyme pausing, and further research should explore the optimization of these amplification parameters in order to quantify the stutter produced for these loci. It must be kept in mind, however that factors which interfere with the pausing of the polymerase may also interfere with the stutter mechanism if it is in deed caused by polymerase pause. This could possibly lead to a lower stutter production as well.

Conclusions

Stutter production is not unique to human identification and the forensic application of STRs. The presence of the additional bands can also lead to difficulty in interpretation and sizing of banding patterns found in research conducted in various fields (i.e. human, plant, and animal research). Although none of the experiments conducted in this project were able to eliminate stutter production during amplification of tetranucleotide STRs, the data suggests that some guidelines may be followed for choosing loci, which minimize its production. Factors influencing stutter production were also identified which can aid in the interpretation of bands whenever stutter may be encountered. In general, a combination of loci and primer sets which produce the

smallest amplicon possible tend to produce lower amounts of stutter than those which give longer amplified product. In addition, the data from this research suggests that tetranucleotide loci with a repeat containing greater than 50% GC may have decreased stutter production (amplification with Taq Stoffel fragment). Further investigation into this result would need to be conducted in order to confirm this phenomenon, due to the inhibition encountered upon amplification of these samples.

APPENDIX I

D5S818 Sequence (GenBank Accession #: G08446)

TCTAATTAAA GTGGTGTCCC AGATAATCTG TACTAATAAA AGTATATTTT
AATAGCAAGT ATGTGACAAG GGTGATTTTC CTCTTTGGTA TCCTTATGTA
ATATTTTGAA GATAGATAGA TAGATAGATA GATAGATAGA TAGATAGATA
GATAGGTAGA TAGAGGTATA AATAAGGATA CAGATATAGN TACAAATGTT
GTAAACTGTG GCTATGATTG GAATCACTTG GCTAAAAAGC GCTNAAGCNT
TCCTCTGNGA GAGGCAATTA CTTTTTTNCT TAGGNACTNC CTCANCAGTC
TNTTNGC



San José State
UNIVERSITY


**Office of the Academic
Vice President**

**Associate Vice President
Graduate Studies and Research**

One Washington Square
San Jose, CA 95192-0025
Voice: 408-924-2480
Fax: 408-924-2477
E-mail: gstudies@wahoo.sjsu.edu
<http://www.sjsu.edu>

35

TO: Cynthia R. Hall
780 N. 18th st.
San Jose, CA 95112

FROM: Craig Cisar, 
Chairperson, Human Subjects Review Board

DATE: March 16, 2000

The Human Subjects-Institutional Review Board has approved
your request to use human subjects in the study entitled:

“Examination of Stutter Produced During
Amplification of Tetranucleotide Short
Tandem Repeat Loci”

This approval is contingent upon the subjects participating in your research project being appropriately protected from risk. This includes the protection of the anonymity of the subjects' identity when they participate in your research project, and with regard to any and all data that may be collected from the subjects. The Board's approval includes continued monitoring of your research by the Board to assure that the subjects are being adequately and properly protected from such risks. If at any time a subject becomes injured or complains of injury, you must notify Nabil Ibrahim, Ph.D., immediately. Injury includes but is not limited to bodily harm, psychological trauma and release of potentially damaging personal information.

Please also be advised that all subjects need to be fully informed and aware that their participation in your research project is voluntary, and that he or she may withdraw from the project at any time. Further, a subject's participation, refusal to participate, or withdrawal will not affect any services the subject is receiving or will receive at the institution in which the research is being conducted.

If you have any questions, please contact me at
(408) 924-2480.

The California State University:
Chancellor's Office
Bakersfield, Chico, Dominguez Hills,
Fresno, Fullerton, Hayward, Humboldt,
Long Beach, Los Angeles, Maritime Academy,
Monterey Bay, Northridge, Pomona,
Sacramento, San Bernardino, San Diego,
San Francisco, San Jose, San Luis Obispo,
San Marcos, Stanislaus, Tulare.

REFERENCES

AmpFISTR Profiler PCR Amplification Kit User's Manual (1997).

AmpFISTR Profiler Plus PCR Amplification Kit User's Manual (1997).

Brewer, Lisa Marie (1995). Frequency Distributions and Linkage Relationships of Ten Genetic Markers in the Santa Clara County Vietnamese Population. Unpublished master's thesis, San Jose State University, California.

Buel, Eric, Schwartz, Margaret B., & LaFountain, Marcia J. (1998). Capillary Electrophoresis STR Analysis: Comparison to Gel-Based Systems. *Journal of Forensic Sciences* 43(1), 164-170.

Clayton, T. M., Whitaker, J. P., Sparkes, R., & Gill, P. (1998). Analysis and Interpretation of Mixed Forensic Stains Using DNA STR Profiling. *Forensic Science International* 91, 55-70.

Demers, Daniel B., Curry, Elizabeth T., Egholm, Michael, & Sozer, Amanda C. (1995). Enhanced PCR Amplification of VNTR Locus D1S80 Using Peptide Nucleic Acid (PNA). *Nucleic Acids Research* 23(15), 3050-3055.

Evetts I W & Gill P. (1991). A Discussion of the Robustness of Methods for Assessing the Evidential Value of DNA Single Locus Profiles in Crime Investigations. *Electrophoresis* 12, 226-230.

Fregeau, Chantal J. & Fournay, Ron M. (1993). DNA Typing with Fluorescently Tagged Short Tandem Repeats: A Sensitive and Accurate Approach to Human Identification. *BioTechniques* 15(1), 100-119.

Gill P, Sullivan K, & Werrett D J, (1990). The Analysis of Hypervariable DNA Profiles: Problems Associated with the Objective Determination of the Probability of a Match. *Hum Gen.* 44, 388-396.

Gill, P., Kimpton, C. P., & Sullivan, K. M. (1992). A Rapid Polymerase Chain Reaction Method for Identifying Fixed Specimens. *Electrophoresis* 13, 173-175.

Huang, Nu En, Schumm, James, & Budowle, Bruce (1995). Chinese Population Data on Three Tetrameric Short Tandem Repeat Loci – HUMTH01, TPOX, and CSF1PO – Derived Using Multiplex PCR and Manual Typing. *Forensic Science International* 71, 131-136.

Kimpton, Colin P., Gill, Peter, Walton, Abbie, Urquhart, Andy, Millican, Emma S., & Adams, Maia (1993). Automated DNA Profiling Employing Multiplex Amplification of Short Tandem Repeat Loci. *PCR Methods Appl.* 3, 13-22.

Lazaruk, K., Walsh, P. S., Oaks, F., Gilbert, D., Rosenblum, B. B., Menchen, S., Scheibler, D., Wenz, H. M., Holt, C., & Wallin, J. (1998). Genotyping of Forensic Short Tandem Repeat (STR) Systems Based on Sizing Precision in a Capillary Electrophoresis Instrument. *Electrophoresis* 19, 86-93.

- Levinson, Gene & Gutman, George A. (1987). Slipped-Strand Mispairing: A Major Mechanism for DNA Sequence Evolution. *Mol Biol Evol.* 4(3), 203-221.
- Lins, Ann M., Sprecher, Cynthia J., Puers, Christoph, & Schumm, James W. (1996). Multiple Sets for the Amplification of Polymorphic Short Tandem Repeat Loci – Silver Stain and Fluorescence Detection. *BioTechniques* 20(5), 882-889.
- Litt, Michael, Hauge, Xueya, & Sharma, Vikram. (1993). Shadow Bands Seen When Typing Polymorphic Dinucleotide Repeats: Some Causes and Cures. *BioTechniques* 15(2), 280-284.
- McDowell, David G., Burns, Nigel A., & Parkes, Helen C. (1998). Localised Sequence Regions Possessing High Melting Temperatures Prevent the Amplification of a DNA Mimic in Competitive PCR. *Nucleic Acids Research* 26(14), 3340-3347.
- Murray, Vincent, Monchawin, Chutima, & England, Phillip R. (1993). The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR. *Nucleic Acids Research* 21(10), 2395-2398.
- Mytelka, Daniel S. & Chamberlin, Michael J. (1996). Analysis and Suppression of DNA Polymerase Pauses Associated with a Trinucleotide Consensus. *Nucleic Acids Research* 24(14), 2774-2781.
- Nolin, Sarah L., Houck, Jr., George E., Gargano, Alice D., Blumstein, Howard, Dobkin, Carl S., & Brown, W. Ted. (1999). FMR1 CGG-Repeat Instability in Single Sperm and Lymphocytes of Fragile-X Premutation Males. *Am J Hum Genet* 65, 680-688.
- Pearson, Christopher E, & Sinden, Richard R. (1996). Alternative Structures in Duplex DNA Formed within the Trinucleotide Repeats of the Myotonic Dystrophy and Fragile X Loci. *Biochemistry* 35, 5041-5053.
- Pearson, Christopher E., Eichler, Evan E., Lorenzetti, Diego, Kramer, Shannon F., Zoghbi, Huda Y., Nelson, David L., & Sinden, Richard R. (1998). Interruptions in the Triplet Repeats of SCA1 and FRAXA Reduce the Propensity and Complexity of Slipped Strand DNA (S-DNA) Formation. *Biochemistry* 37, 2701-2708.
- Reynolds, Rebecca, Sensabaugh, George, & Blake, Edward. (1990). Analysis of Genetic Markers in Forensic DNA Samples Using the Polymerase Chain Reaction. *Analytical Chemistry* 63(1).
- Roy, Reena, Steffens, David, Gartside, Bill, Jang, Gi Y., & Brumbaugh, John A. (1996). Producing STR Locus Patterns from Bloodstains and Other Forensic Samples Using an Infrared Fluorescent Automated DNA Sequencer. *Journal of Forensic Sciences* 41(3), 418-424.
- Sarkar, Gobinda, Kapelner, Stephen, & Sommer, Steve S. (1990). Formamide Can Dramatically Improve the Specificity of PCR. *Nucleic Acids Research* 18(24), 7465.
- Schlotterer, Christian & Tautz, Diethard, (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research* 20(2), 211-215.

Smith, R. N. (1995). Accurate Size Comparison of Short Tandem Repeat Alleles Amplified by PCR. *BioTechniques* 18(1), 122-128.

Sprecher, Cynthia J., Puers, Christopher, Lins, Ann M., & Schumm, James W. (1996). General Approach to Analysis of Polymorphic Short Tandem Repeat Loci. *BioTechniques* 20(2), 266-276.

Walsh, P. S., Erlich, H. A., & Higuchi, R. (1992). Preferential PCR Amplification of Alleles: Mechanisms and Solutions. *PCR Methods Appl.* 1, 241-250.

Walsh, P. Sean, Fildes, Nicola J., & Reynolds, Rebecca. (1996). Sequence Analysis and Characterization of Stutter Products at the Tetranucleotide Repeat Locus vWA. *Nucleic Acids Research* 24(14), 2807-2812.

Whitaker, J. P., Clayton, T. M., Urquhart, A. J., Millican, E. S., Downes, T. J., Kimpton, C. P., & Gill, P. (1995). Short Tandem Repeat Typing of Bodies from a Mass Disaster: High Success Rate and Characteristic Amplification Patterns in Highly Degraded Samples. *BioTechniques* 18(4), 670-677.

White, Peter J., Borts, Rhona H., & Hirst, Mark C. (1999). Stability of the Human Fragile X (CGG)_n Triplet Repeat Array in *Saccharomyces cerevisiae* Deficient in Aspects of DNA Metabolism. *Mol Cell Bio* 19 5675-5684.